Heuristics as Decision Rules*
— Part I: The Single Consumer —

by

Werner Güth† and Wilhelm Neuefeind‡

Abstract: Many consumption prices are highly volatile. It would certainly overburden our cognitive system to fully adjust to all these changes. Households therefore often rely on simple heuristics when deciding what to consume, e.g. in the form of a constant budget share for a specific consumption commodity, like a vacation, or of a constant consumption amount for low-cost commodities as food items. Using utility functions we can measure the welfare loss, caused by such heuristics, and to what extent this can be reduced by adaptation. In the present Part I the analysis is mainly restricted to a single consumer with a Cobb-Douglas utility function. General utility functions will also be considered. Part II will study exchange economies.

1 *Introduction*

Traditional microeconomic theory assumes that consumers optimally adjust to prices which are simultaneously determined through market clearing. This has led to a well developed theory of household behavior and trade (see for instance Mas-Colell, Whinston & Green (1993)), which, however, is not based on a realistic model of consumers' behavior. Here we do not assume immediate and perfect adjustment to prices. Consumers may not even have the latest information on prices when determining their consumption behavior. This will, of course, often render their consumption behavior suboptimal. On the other hand consumers are by no means neglecting new information on prices; they rather frequently try to improve their choice in the light of earlier experiences. In such a setup one can ask

1. whether the results depend on the type of choice variable (in our study expenditure shares or consumption amounts) on which consumers rely,
2. how the parameters of the adaptation dynamics influence the welfare losses and the process,

---

*This research was conducted while both authors were visiting the Zentrum für interdisziplinäre Forschung an der Universität Bielefeld. The visits were supported by the Deutsche Forschungsgemeinschaft (DFG) via SFB 373. We acknowledge helpful comments by Rainer Schulz and Reinhard Selten.

†Humboldt University Berlin

‡Washington University in St. Louis

3. whether the welfare losses due to relying on learning heuristics rather than on optimization decrease or even vanish over time,
4. how the stable distributions of the resulting stochastic process depend on the parameters of the dynamics, and
5. whether and how the interaction of dynamically adjusting agents in a general equilibrium model modifies the conclusions.

Part I of our study concentrates on (1), (2), (3), and (4) whereas Part II will explore question (5).

More specifically, it will be assumed that success in the sense of *ex post*-satisfaction, which determines the adaptation dynamics, is measured by a utility function. Regardless which degree of happiness the consumer experiences (as measured by the utility function), she does not attempt to maximize her utility. She rather gradually adjusts her choice variable as somehow suggested by past experience. As choice variables of the consumer we investigate

(a) consumption expenditures and

(z) consumption amounts.

Due to the choice of the specific Cobb-Douglas utility function, in the case of stochastic price expectations, an optimal adjustment is possible for a, whereas z-heuristics excludes perfect adjustment.

The adaptation dynamics is close to directional learning. A consumer, who has changed her decision, later learns whether this has increased or decreased her satisfaction without being aware how exactly these changes depend on her choice variable and on the (deterministic or stochastic) price level. If the former change has been successful, the consumer will continue to move in the same direction whereas otherwise she will redo the former change. By introducing a smallest grid and assuming that all changes be local, i.e. of just one step on the grid, this completely determines the adaptation where, of course, the adjustment process is stochastic in case of stochastic prices. We, however, will introduce a stochastic element even in the case of fixed prices.

Usual adaptation dynamics assume that agents may try out radically different strategies rather than experiment with directional changes. Reinforcement learning (Bush-Mosteller (1955) and Roth-Erev (1998)), for instance, assumes that agents use their different behaviors with probabilities, determined by their respective previous accumulated success. Best reply dynamics or variants thereof are essentially belief dynamics which do not question decision rationality with respect to prevailing beliefs. They assume awareness of the true success function and forward looking rational deliberation (see Fudenberg & Levine (1998)) which we wanted to substitute by decision heuristics and adaptation to past experiences.

Evolutionary dynamics (see Hammerstein & Selten (1994) and Weibull (1995) for surveys) deny any cognition by just focusing on the survival prospects of possible strategies. More specifically, one asks which behavior is best or evolutionarily stable if reproductive success is measured by the true utility. When considering isolated consumers, the stable result is obviously maximizing the true utility over the set of all possible realisations of the variables of the heuristics. Although other interpretations are possible, too, the typical interpretation would be the one of evolutionary biology: A more successful consumer has more offspring so that over time more and more consumers would rely on the best variant of the possible decision heuristic.

In contrast to those concepts, directional learning (for applications and experimental support see Selten & Buchta (1999)) requires a cognitive representation of the decision environment by the decision maker. It assumes that she, after receiving feedback information about her former choice, e.g. by experiencing her *ex post*-satisfaction as measured by the utility function, asks herself whether it would have been better to have chosen a larger or smaller activity level. If so, she either changes her behaviour in the desired direction or not at all. Thus what is excluded is only a change in the non-desired direction. How far she moves in the desired direction, if she does, is not at all specified, i.e. it is a qualitative theory of direction learning.

Our adaptation dynamics also looks at the direction of changes, but based on past experiences with such changes rather than based on *ex post*-valuations (assuming a correct cognitive representation of the decision environment). Thus the main difference of our adaptive dynamics and directional learning seems to be that we substitute awareness of the decision environment by past experiences with directional adaptation. Other differences are minor: If a change has been successful, she will continue to move in this direction only with a given adaptation rate. If she has not recently experimented with behavioral changes, she will mutate locally in either direction. The purely qualitative aspect is finally avoided by assuming a smallest grid and only local changes of just one step in a grid.

A similar study, regarding the basic methodology, is Huck, Normann & Oechsler (1998), whose adaptive dynamics correspond very closely to ours in the sense that agents continue to move in one direction if this has been successful before. The study of homogeneous oligopoly markets, however, denies any stochastic aspects of the decision environment which we consider as crucial when judging the reliability of certain decision heuristics. These authors also explore different action variables, namely sales amounts and sales prices, and a different environment. These somewhat restrictive assumptions , however, allow them to derive a very interesting result, namely that the stable result is cooperation. It will be interesting to see whether a similar conclusion results for the simple exchange economy.

The report will be organized as follows: In Section 2, we lay out our model of a single consumer confronting a deterministic price path. Section 3 illustrates aspects and properties of the adaptive process which is theoretically explored in Section 4. Section 5 formally introduces price uncertainty. Section 6 studies the effect of price uncertainty via simulations. Section 7 concludes. In Part II of this project we plan to analyse the interactive adaptation dynamics of several interacting consumers in deterministic and stochastic exchange economies.

2 The Household Model

Since we focus on consumption heuristics for a specific consumption good, we assume just two commodities x and y, interpreting commodity y as "all other consumption goods." Choosing commodity y as numeraire commodity, i.e. setting its price equal to 1, we just have one variable price p, the price of commodity x. While we are especially interested in a randomly fluctuating price, for the sake of exposition we will present the case of a fixed price first. To save on notation we assume a constant disposable income of 1 which is completely reserved for consumption purposes. To measure the loss, implied by relying on periodic heuristics, and to provide some success measure when judging whether a previous change has been good or bad *ex post*-satisfaction is assumed to be determined by a utility function $u(x, y)$. Much of the analysis holds for any non-satiated, continuous, strictly quasi-concave utility function[1].

For the sake of exposition, however, we will introduce the framework via a Cobb-Douglas utility function

$$u(x, y) = x^{\alpha} \cdot y^{1-\alpha} \quad \text{with} \quad 0 < \alpha < 1. \tag{1}$$

Here $x, y \geq 0$ denote the consumption amounts of the two commodities

[1]For an introduction and discussion of these properties see Mas-Colell, Whinston & Green (1993).

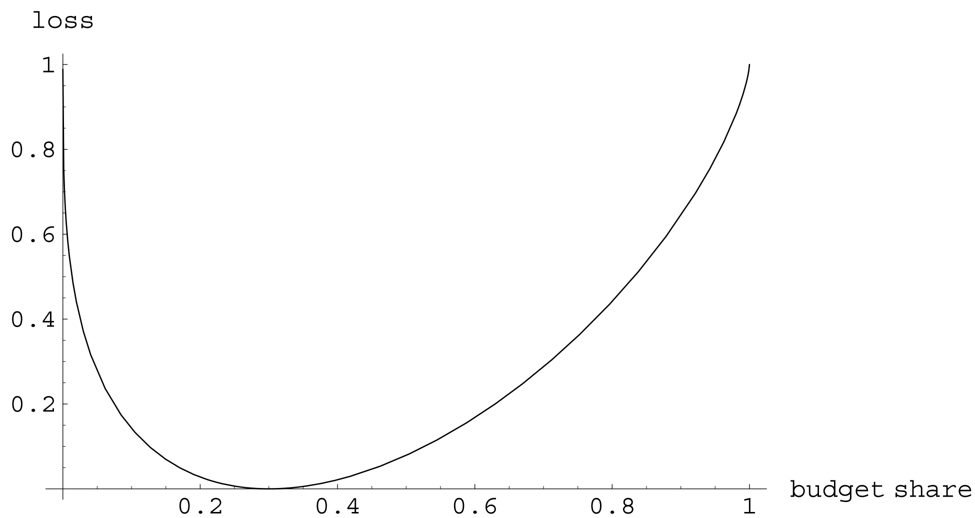

Figure 1: Welfare Loss for a-heuristics with $\alpha = .3$

and $u(x, y)$ the satisfaction, generated by the consumption vector (x, y). Our idea is that the consumer may be quite uncertain how $u(\cdot)$ depends on her choices. In the spirit of boundedly rational behavior the function $u(\cdot)$ only guides her when judging whether or not previous behavioral changes were encouraging or not. Fully rational behavior according to classical consumer's theory for this utility function $u(\cdot)$ would require

$$x^*(p) = \frac{\alpha}{p} \quad \text{and} \quad y^*(p) = 1 - \alpha, \tag{2}$$

and yield a utility level of

$$(\frac{\alpha}{p})^{\alpha}(1 - \alpha)^{1-\alpha}. \tag{3}$$

Clearly, every price p would require a specific amount $x^*(p)$, i.e. the household would have to adjust to any change in the price p.

Our approach is that the household, being more or less unaware of $u(\cdot)$, does not optimize, but rather relies on decision heuristics, e.g. by spending a chosen share a with $0 < a < 1$ of the unit income for the consumption of commodity 1 or by consuming a chosen amount z of commodity 1 satisfying $0 \leq z$ and $zp_u < 1$, where p_u is the upper bound of the price p.

In the case of a constant budget share a, commodity 1 could be the percentage of a family's disposable income set aside for the yearly vacation. The quantity of commodity 1 would then measure the length of the recreation or (with appropriate interpretation) the quality of the holiday resort chosen. Constant consumption amounts z are rather typical for food

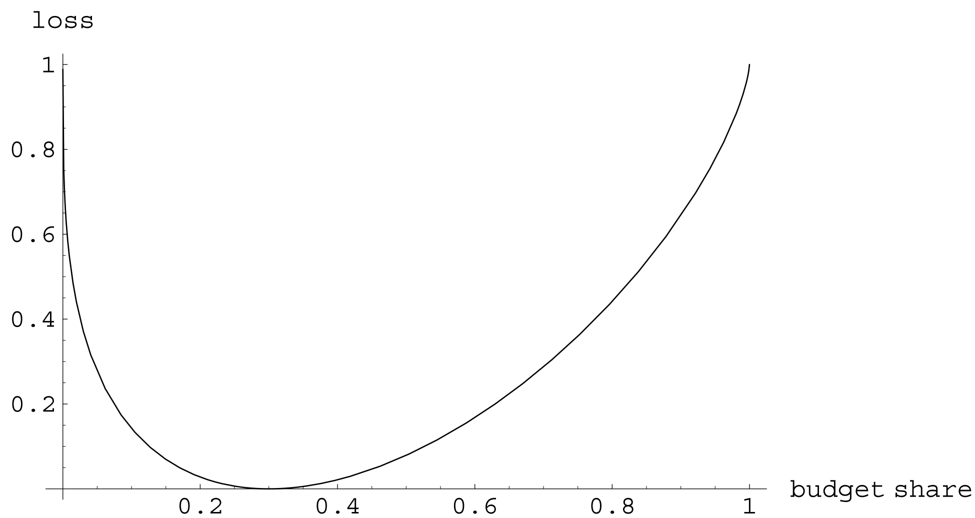

Figure 2: Welfare Loss for z-heuristics with $\alpha = .3$ and $p = .5$

products for which the assumption $zp_u < 1$ imposes hardly any restriction. We will call the first approach a-heuristics/-dynamics while the latter will be referred to as z-heuristics/-dynamics. Now the a-heuristics yield the purchased quantities

$$x(a,p) \quad = \quad \frac{a}{p} \quad \text{and} \quad y(a,p) = 1 - a \tag{4}$$

with a utility level, depending on a, of

$$V(a) = (\frac{a}{p})^{\alpha}(1-a)^{1-\alpha}. \tag{5}$$

We will use the term "indirect utility" for V. Thus the relative loss in well-being over the whole period is measured by

$$L(a) = 1 - V(a)/V(\alpha) = 1 - (\frac{a}{p})^{\alpha}(1-a)^{1-\alpha}/(\frac{\alpha}{p})^{\alpha}(1-\alpha)^{1-\alpha}. \tag{6}$$

For the z-heuristics the corresponding purchased quantities are

$$x(z,p) \quad = \quad z \quad \text{and} \quad y(z,p) = 1 - pz. \tag{7}$$

The corresponding indirect utility will be denoted W and the relative loss in well-being is

$$M(z) = 1 - z^{\alpha}(1-pz)^{1-\alpha}/(\frac{\alpha}{p})^{\alpha}(1-\alpha)^{1-\alpha}. \tag{8}$$

From Figure 1 and Figure 2 it becomes clear that the choice of $a = \alpha$ reduces the welfare loss to zero for **all** prices p. For the z-heuristics, the choice of $z = \alpha/p$ does the same trick for the fixed price p, but we would expect persistent welfare losses if prices change. As long as the agent's utility function shows standard features (as being derived from a strictly convex preference relation) both heuristics will lead to optimal behavior.[2] Thus the loss in well-being would get smaller if a or z converges over time to α or α/p, respectively since $L(\alpha) = M(\alpha/p) = 0$. For the special case of the Cobb-Douglas utility function with $\alpha = .3$, $p = .5$, Figure 1 offers the graph of $L(a)$, while Figure 2 presents the graph of $M(z)$. Note that in Figure 1 the function attains its minumum at the argument α, while in Figure 2 the argument, at which the minumum is attained, equals α/p, thus it depends on p.

3 Adjusting to Past Success

For the a- and z-heuristics let a_t and z_t, respectively, denote the value chosen in period t. Adaptation takes the form of changing a_t or z_t, respectively, in the light of earlier experiences. At time t, when our consumer has to choose a_t, she recalls a_{t-1} and a_{t-2}, which also allows her to recognize her well-being in the corresponding time periods, and consequently the change thereof. We have thus restricted ourselves to the most recent experiences. We find it convenient to introduce

$$\Delta_t a \quad = \quad a_t - a_{t-1} \tag{9}$$

as the a-change from period $t-1$ to period t and similarly

$$\Delta_t z \quad = \quad z_t - z_{t-1}. \tag{10}$$

The state (of knowledge) of the consumer at time t is given by the pair $(a_{t-1}, \Delta_{t-1}a)$. How does a household judge whether a previous adjustment $\Delta_{t-1}a$ or $\Delta_{t-1}z$ has been successful or not? The consumer will react to the change in indirect utility as caused by her previous change:[3]

$$\Delta V(a_{t-1}, \Delta_{t-1}a) \quad = \quad V(a_{t-1}) - V(a_{t-2}) \quad \text{and} \tag{11}$$

$$\Delta W(z_{t-1}, \Delta_{t-1}z) \quad = \quad W(z_{t-1}) - W(z_{t-2}), \text{ respectively.} \tag{12}$$

Adaptation is understood here as systematic and intuitive behavioral changes in the light of previous experiences with such changes. When adapting behavior we consumers usually do not jump from one extreme to

[2] subject to minor qualifications

[3] Because p is fixed here, we suppress the dependence of ΔV and ΔW on p.

another. We mimic this real life behavior by restricting adaptations to the choice of a neighboring value. Let g (a positive number) denote the size of this smallest step. Thus

$$a \in \mathcal{G} \quad = \quad \{g, 2g, \ldots\ Gg\} \quad \text{with} \quad Gg < 1. \tag{13}$$

In the same way let denote h the step size for z-changes, i.e.

$$z \in \mathcal{H} \quad = \quad \{h, 2h, \ldots\ Hh\} \quad \text{with} \quad Hh < p_u^{-1}. \tag{14}$$

If $\Delta_{t-1}a = 0$ or $\Delta_{t-1}z = 0$, one has not recently experimented with behavioral changes so that adaptation in the light of recent experiences is impossible. Similarly, when $\Delta_{t-1}a \neq 0$, but $\Delta V(a_{t-1}, \Delta_{t-1}a) = 0$ or $\Delta_{t-1}z \neq 0$, but $\Delta W(z_{t-1}, \Delta_{t-1}z) = 0$, experimentation does not provide a clue whether to change further in the previous direction or to move back. In order to prevent that learning then stops, we assume that households switch to a neighboring parameter value with small, but positive mutation probability ε, where in the literature one is typically interested in limit results for $\varepsilon \to 0$. We will put the emphasis on small but finite ε, but will discuss limit behavior as well. Formally, it is assumed that in case of

$$\Delta V(a_{t-1}, \Delta_{t-1}a) \cdot \Delta_{t-1}a \quad = \quad 0 \tag{15}$$

mutation is governed by ε as follows:

$$\Delta_t a \quad = \quad (+g|\frac{\varepsilon}{2},\ 0|1-\varepsilon,\ -g|\frac{\varepsilon}{2}). \tag{16}$$

The notational convention here is that our consumer "mutates" up or down by g with probability $\varepsilon/2$ each and keeps the old level of a with probability $1 - \varepsilon$. In the same way for

$$\Delta W(z_{t-1}, \Delta_{t-1}z) \cdot \Delta_{t-1}z \quad = \quad 0 \tag{17}$$

we assume

$$\Delta_t z \quad = \quad (+h|\frac{\varepsilon}{2},\ 0|1-\varepsilon,\ -h|\frac{\varepsilon}{2}). \tag{18}$$

If (15) respectively (17), does not hold, experimentation has been informative. Recall that even in this case, the consumer adapts by a rate short of 1. Let δ with $0 < \delta < 1$ denote this adaptation rate, i.e. the typically (compared to ε) large probability of adapting in the light of one's recent experiences. It is an obvious idea in the case of a positive ($\Delta V(\cdot,\cdot) > 0$ or $\Delta W(\cdot,\cdot) > 0$) recent change in well-being to move further in the previously experienced direction. Similarly, in case of a negative ($\Delta V(\cdot,\cdot) < 0$

or $\Delta W(\cdot,\cdot) < 0$) recent change, one will want to switch back, i.e. to choose the previous parameter value $a_t = a_{t-2}$ and $z_t = z_{t-2}$, respectively. More formally,

$$\Delta_t a = \begin{cases} (+g|\delta,\ 0|1-\delta) & \text{if} \quad \Delta V(a_{t-1}, \Delta_{t-1}a)\cdot\Delta_{t-1}a > 0 \\ (-g|\delta,\ 0|1-\delta) & \text{if} \quad \Delta V(a_{t-1}, \Delta_{t-1}a)\cdot\Delta_{t-1}a < 0 \end{cases} \tag{19}$$

and

$$\Delta_t z = \begin{cases} (+h|\delta,\ 0|1-\delta) & \text{if} \quad \Delta W(z_{t-1}, \Delta_{t-1}z)\cdot\Delta_{t-1}z \quad > 0 \\ (-h|\delta,\ 0|1-\delta) & \text{if} \quad \Delta W(z_{t-1}, \Delta_{t-1}z)\cdot\Delta_{t-1}z \quad < 0. \end{cases} \tag{20}$$

Of course, in a world, in which a consumer cannot sell short of any commodity, the non-negativity of a, and $1-a$, or z and $1-pz$, respectively, requires some changes for the border levels of a and z. In our simulations of the a-heuristics, however, we prefer to rely on the self-healing aspect of the adaptation dynamics: whenever a or $1-a$ becomes negative, the consumer's well-being suffers so much that she automatically increases the corresponding number sooner or later. For the z-heuristics this approach turns out to be somewhat problematic in the case of stochastic prices. Therefore we will choose the obvious modification of not allowing the consumer to choose a z being negative or larger than $1/p_u$. This completely describes the dynamics of a_t and z_t, respectively, as influenced by mutation (where the switching probability ε should be small) and by adaptation (where the switching probability δ should be large). In the following we want to study both,

- the long run-results of such behavioral dynamics, and
- the processes leading to them

to assess the losses in well-being by not adjusting or by adjusting too slowly.

4 Dynamic Analysis

Whether our consumer's starting situation is certain or a probability distribution, in all subsequent time periods, her situation can only be described by a vector of probabilities over all possible states $(a_t, \Delta_t a)$. For $(a_t, \Delta_t a)$ let $r(a_t, \Delta_t a)$ denote the probability of being in that state. Similarly $s(z_t, \Delta_t z)$ is the probability of being in state $(z_t, \Delta_t z)$. The probability of all possible states $(a_t, \Delta_t a)$, respectively $(z_t, \Delta_t z)$, can be described by vectors with $2+3(G-2)+2 = (3G-2)$ and $3H-2$ components, respectively:

$$r(t) \quad = \quad (r(a_t, \Delta_t a))_{\text{all states } (a_t, \Delta_t a)}, \text{ and} \tag{21}$$

$$s(t) \quad = \quad (s(z_t, \Delta_t z))_{\text{all states } (z_t, \Delta_t z)}. \tag{22}$$

Now (13), (15), (16), (19), and (21) imply a stochastic process which can be described with help of a $(3G-2)$ by $(3G-2)$ transition matrix R; i.e. our process is a Markov chain. The probabilities $r(t)$ then are generated from $r(t-1)$ by

$$r(t) \quad = \quad r(t-1)R. \tag{23}$$

Similarly, (14), (17), (18), (20), and (22) determine the stochastic dynamics

$$s(t) \quad = \quad s(t-1)S \tag{24}$$

where S is a $(3H-2)$ by $(3H-2)$ transition matrix.

Important information on the dynamics of the a- and z-heuristics, respectively, is given by the existence and properties of the corresponding stationary vectors r and s satisfying

$$r \quad = \quad rR \quad \text{and} \tag{25}$$

$$s \quad = \quad sS, \quad \text{respectively.} \tag{26}$$

Such a vector is collinear to a stationary distribution and may have to be normalized before its components add up to 1 and, thus, is a stationary distribution.

For such an analytic treatment of the behavioral model we will take care of the boundary cases in the manner alluded to earlier for the z-heuristics. It is simply assumed that she mutates in the only possible direction with full mutation probability ε and that an impossible adaptation step does not take place.

It is not difficult to compute the transition matrix for any of the cases included in our analysis. It consists of G small block matrices (their dimensions not exceeding 3×5) on the diagonal. For illustration, we give the matrix for the least complicated case $G = 3$. Let $c = .5$ and $g = .25$. With $\mathcal{G} = \{c-g, c, c+g\}$ symmetrically written, our state space is, conveniently ordered:

$$\{(c-g, 0), (c-g, -g), (c, g), (c, 0), (c, -g), (c+g, g), (c+g, 0)\}.$$

The 7 x 7 transition matrix R then is given as follows:

$$\begin{pmatrix} 1-\varepsilon & 0 & \varepsilon & 0 & 0 & 0 & 0 \\ 1-\delta P_{(c-g,-g)} & 0 & \delta P_{(c-g,-g)} & 0 & 0 & 0 & 0 \\ 0 & \delta P_{(c,g)} & 0 & 1-\delta & 0 & \delta P^{(c,g)} & 0 \\ 0 & \varepsilon/2 & 0 & 1-\varepsilon & 0 & \varepsilon/2 & 0 \\ 0 & \delta P^{(c,-g)} & 0 & 1-\delta & 0 & \delta P_{(c,-g)} & 0 \\ 0 & 0 & 0 & 0 & \delta P_{(c+g,g)} & 0 & 1-\delta P_{(c+g,g)} \\ 0 & 0 & 0 & 0 & \varepsilon & 0 & 1-\varepsilon \end{pmatrix}$$

The zero entries are evident from the definition of the process as well as the entries concerning $\varepsilon-$mutation in rows 1, 4, and 7. The entries in the other rows require some explanation. The symbol $P_{(c-g,-g)}$ denotes the probability that a *previous* move from $a_{t-2} = c - g - (-g) = c$ to $a_{t-1} = c - g$ led to $V(c-g) < V(c)$; i.e. a **lower** well-being, thus the subscript. In this case, the previous move would be reversed; i.e $\Delta_t a = g$ and $a_t = c - g + g = c$. For the other superscripted P's which apply in case of a **higher** well-being, the meaning of the symbol $P^{(\cdot,\cdot)}$ is clear, too.

In the case of fixed prices, we are currently presenting, the probabilities $P_{(\cdot,\cdot)}$ and $P^{(\cdot,\cdot)}$ are degenerate: they only take on the values 1 and 0. In the case of price uncertainty, which will be presented in the next section, these probabilities are non-degenerate. The derivation of the probabilities is somewhat technical and will be relegated to an Appendix.

There are some standard theoretical results about Markov chains (see e.g. Kemeny-Snell (1960)) which indicate that the sequence of state vectors $r(t)$ from (23) converges. For any combination of (ε, δ) with $\varepsilon > 0$ and $\delta < 1$, a stationary distribution r (see (25)) exists. The distributions are unique. For $\varepsilon = 1-\delta = 0$ the distributions also exist but are not necessarily unique.

Due to the simple structure of the transition matrix, say R, the stationary distributions can be explicitly computed – even for general utility functions and more general state spaces.[4] While the points in $\mathcal{G}$ do not have to be equi-distant for this computation to be valid, we will consider our special $\mathcal{G}$ – for notational convenience only.

Strictly speaking, six different cases would have to be considered. To see this, note that a maximum can be interior or one of two boundary maxima. Recall that our assumptions on the consumer's utility function u assure a unique maximum on any budget line generated by positive prices. Denote by v_i the utility of the consumption bundle generated by the choice of ig as budget share. Then $(v_i, i = 1, \ldots, G)$ has not more than two maxima. The cases involving two maxima are non-generic and can be

[4] We thank Reinhardt Selten whose insistence that this could be done for more general functions induced us to derive these results.

omitted without loss of generality. Of the remaining three cases, the case of the unique "interior" maximum is the most important. We will present this case in some detail. The other two cases of one "boundary" maximum are less important. In addition, the modifications of the computations and formulae are straightforward. Thus these cases will be omitted, too.

Let $i := \arg\max_{j=1}^{G} v_j$. Since the maximum is interior, $1 < i < G$. Because $\mathcal{G}$ (almost) is a product space, we can use a convenient notation for the entries of the stationary vector r: for any $i = 1, \ldots, G$, let $r_{i,k}$ be the probability (possibly after normalisation) of being in state $(ig, kg) \in \mathcal{S}$ with $k \in \{-1, 0, +1\}$. Note that the two index pairs $1, 1$ and $G, -1$ are excluded.

We normalize $r_{i,0}$ to equal 1. Then, for any $\varepsilon > 0$ and $\delta < 1$, we recursively compute the other entries of r "from the inside out." The following result holds whether $i = G/2$ or not.

Result: The unique stationary distribution for R is collinear with r where r is computed as follows

1. $r_{i,0} = 1$.
2. $r_{i,+1} = r_{i,-1} = \varepsilon/(2(1-\delta))$.
3. For j in the *increasing* utility range $0 < j < i$,
$$r_{j,k} = r_{j+1,+1} \cdot \begin{cases} 1 & \text{for} \quad k = -1 \\ 2(1-\delta)/(\varepsilon(1+\delta)) & \text{for} \quad k = 0 \\ (1-\delta)/(1+\delta) & \text{for} \quad k = +1. \end{cases}$$
4. And symmetrically, for j in the *decreasing* utility range $i < j < G$,
$$r_{j,k} = r_{j-1,-1} \cdot \begin{cases} 1 & \text{for} \quad k = +1 \\ 2(1-\delta)/(\varepsilon(1+\delta)) & \text{for} \quad k = 0 \\ (1-\delta)/(1+\delta) & \text{for} \quad k = -1. \end{cases}$$
5. Finally, we obtain
$$r_{1,k} = r_{2,+1} \cdot \begin{cases} (1-\delta)/\varepsilon & \text{for} \quad k = 0 \\ 1 & \text{for} \quad k = -1. \end{cases}$$
and
$$r_{G,k} = r_{G-1,-1} \cdot \begin{cases} 1 & \text{for} \quad k = +1 \\ (1-\delta)/\varepsilon & \text{for} \quad k = 0. \end{cases}$$

The proof of this statement simply consists of verifying the equality $r = rR$. Because the recursive definition of r makes this verification a cumbersome task for general or large G, we will omit this verification. The reader, however, will find it easy to verify it for the case $G = 3$. r is collinear to the distribution $(4(1 - \delta + \varepsilon))^{-1}(1-\delta, \varepsilon, \varepsilon, 2(1-\delta), \varepsilon, \varepsilon, 1-\delta)$. Because there is no price uncertainty, $P_{(c-g,-g)} = P_{(c+g,g)} = P^{(c,g)} = P^{(c,-g)} = 1$, and the two other probabilities $P_{(\cdot,\cdot)}$ both are zero. Thus the matrix R simplifies to

$$\begin{pmatrix} 1-\varepsilon & 0 & \varepsilon & 0 & 0 & 0 & 0 \\ 1-\delta & 0 & \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\delta & 0 & \delta & 0 \\ 0 & \varepsilon/2 & 0 & 1-\varepsilon & 0 & \varepsilon/2 & 0 \\ 0 & \delta & 0 & 1-\delta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta & 0 & 1-\delta \\ 0 & 0 & 0 & 0 & \varepsilon & 0 & 1-\varepsilon \end{pmatrix}.$$

The verification of equation (25) now is straightforward.

The economically more relevant information consists of the probabilities of ending up in particular states in $\mathcal{G}$ as opposed in $\mathcal{S}$. We thus collaps the state space to its natural old self $\mathcal{G} = \{c-g, c, c+g\}$ and compute the $G-$dimensional vector of probabilities q by adding up the probabilities in r belonging to each state in $\mathcal{G}$. For our example this is simple and yields the vector $(0.25, 0.5, 0.25)$ which seems surprising because it does not depend on ε or δ. Upon reflection however, one notices that the data which enter the process are in fact symmetric and independent of ε and δ.

For the general situation, the task is tedious but not particularly difficult. The result is that the stable distribution on $\mathcal{G}$ is collinear with the vector q defined relative to the maximizing point i:

For $j = 1, \cdots, G$, we obtain as the "probability" to be in state $i + j$,

$$q_{i+j} = \begin{cases} 1 & \text{for} \quad j = 0 \\ ((1-\delta)/(1+\delta))^{|j|-1}/(1-\delta) & \text{for} \quad j \in \{-i+2, \cdots, -1, \\ & \qquad\quad +1, \cdots, G-i-2\} \\ ((1-\delta)/(1+\delta))^{|j|-1}/2 & \text{for} \quad j \in \{-i+1, G-i\}. \end{cases}$$

Note that for the "interior" maximum case the stable distribution on $\mathcal{G}$ is independent of ε. For the "boundary" maximum cases, this is no longer true as can be seen from Table 4.2 and Table 4.3.

We now compute the stationary distribution for various combinations of the mutation and adaptation probabilities for the Cobb-Douglas utility function. The results, collected in the following tables, illustrate how the stable distributions r in the sence of (25) depend on the interaction of the two parameters ε and δ. For illustration only, in Table 4.1 we give the stable distributions for $\alpha = .5$ and $\delta = .9$, for two values of ε. Note that the numbers in Table 4.1 are consistent with $q = (0.25, 0.5, 0.25)$ independently of ε.

	$\delta = 0.9$
$\varepsilon = 0.1$	$(0.125, 0.125, 0.125, 0.250, 0.125, 0.125, 0.125)$
$\varepsilon = 0.00001$	$(.249975, .000025, .000025, .49995, .000025, .000025, .249975)$

Table 4.1: An Example of Stable Distributions for $\alpha = 0.5$

For the "boundary" maximum case $\alpha = 0.2$, slightly different computation formulae (which have been omitted) have to be used. The results are somewhat more interesting, as indicated in Table 4.2.

$\delta = 0.9$	$\varepsilon = 0.1$	$(0.8326693, 0.1593625, 0.0079681)$
	$\varepsilon = 0.0001$	$(0.9003933, 0.0948635, 0.0047432)$
$\delta = 0.9999$	$\varepsilon = 0.1$	$(0.9165826, 0.0834132, 0.0000042)$
	$\varepsilon = 0.0001$	$(0.9999800, 0.0000200, 0.0000000)$

Table 4.2: Stable Distributions in $\mathcal{G}$ for $\alpha = 0.2$

It is interesting to note the lack of continuity at $\varepsilon = 0$ and $\delta = 1$, for which the set of stable distributions is $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

For fixed $\delta < 1$ and $\varepsilon \to 0$ the stable distributions converge to $(1, 0, 0)$, a fact which can be corroborated by consulting Table 4.2. For fixed $\varepsilon > 0$ an $\delta \to 1$ the corresponding stable distributions converge to a distribution s_ε, which depends on ε. Such results are collected in Table 4.3.

	s_ε
$\varepsilon = 10^{-1}$	$(0.9166667, 0.0833333, 0.0000000)$
$\varepsilon = 10^{-5}$	$(0.9999900, 0.0000100, 0.0000000)$
$\varepsilon = 10^{-9}$	$(1.0000000, 0.0000000, 0.0000000)$

Table 4.3: Limits of Stable Distributions for $\alpha = 0.2$ and $\delta \to 1$

When analyzing the welfare consequences by computing the expected utility of each of the stable distributions, one discovers that welfare grows with δ and $1-\varepsilon$ but can be much lower if one happens to end up in one of the "wrong" stable distributions for $\varepsilon = 1-\delta = 0$.

We conclude this section by giving an indication to what extent the use of the Cobb-Douglas utility function can ease an analysis. In the special case of the a-heuristics with fixed prices, we obtain the following identities.

$$\begin{array}{llll} P_{(c-g,-g)} & = P^{(c,g)}; & P_{(c,-g)} & = 1 - P^{(c,-g)}; \\ P_{(c,g)} & = 1 - P^{(c,g)}; & P_{(c+g,g)} & = P^{(c,-g)}. \end{array}$$

Using these substitutions, we obtain a matrix of transition probabilities, whose entries depend on only **two** parameters possibly varying with α, as follows:

$$\begin{pmatrix} 1-\varepsilon & 0 & \varepsilon & 0 & 0 & 0 & 0 \\ 1-\delta P^{(c,g)} & 0 & \delta P^{(c,g)} & 0 & 0 & 0 & 0 \\ 0 & \delta(1-P^{(c,g)}) & 0 & 1-\delta & 0 & \delta P^{(c,g)} & 0 \\ 0 & \varepsilon/2 & 0 & 1-\varepsilon & 0 & \varepsilon/2 & 0 \\ 0 & \delta P^{(c,-g)} & 0 & 1-\delta & 0 & \delta(1-P^{(c,-g)}) & 0 \\ 0 & 0 & 0 & 0 & \delta P^{(c,-g)} & 0 & 1-\delta P^{(c,-g)} \\ 0 & 0 & 0 & 0 & \varepsilon & 0 & 1-\varepsilon \end{pmatrix}$$

The study of the influence of changes in the setting on the behavior of the consumer would thus be greatly facilitated.

5 *Price Uncertainty*

A source of consumers' inability or unwillingness to optimally adjust budget shares or consumption quantities to changing prices lies in the fact that prices often rapidly and/or randomly change. We will model this volatility in two different ways. First we consider the case that prices change randomly and frequently in any period between adjustments of budget shares or quantities and that the consumer buys frequently and is thus exposed to a large representative sample of prices. If the distribution of prices is given by a density $\psi : R_{++} \to R_+$, it is reasonable to measure the consumer's well-being as the average well-being over all prices

$$\int (\frac{a}{p})^\alpha \cdot (1-a)^{1-\alpha}\psi(p)dp, \text{ and} \tag{27}$$

$$\int z^\alpha (1-zp)^{1-\alpha}\psi(p)dp, \text{ respectively.} \tag{28}$$

In the case of the a-heuristics we obtain:

$$V(a) \quad = \quad a^{\alpha}(1-a)^{1-\alpha}\int(\frac{1}{p})^{\alpha}\psi(p)dp \quad \text{and} \tag{29}$$

$$L(a) = 1 - a^{\alpha}(1-a)^{1-\alpha}/\alpha^{\alpha}(1-\alpha)^{1-\alpha}. \tag{30}$$

For the case of the z-heuristics the expressions are not quite as simple but still straightforward:

$$W(z) \quad = \quad z^{\alpha}\int(1-zp)^{1-\alpha}\psi(p)dp \quad \text{and} \tag{31}$$

$$M(z) = 1 - z^{\alpha}\int(1-zp)^{1-\alpha}\psi(p)dp/(\alpha^{\alpha}(1-\alpha)^{1-\alpha}\int(\frac{1}{p})^{\alpha}\psi(p)dp). \tag{32}$$

For a wide variety of densities, in particular, for the uniform density between a minimal value p_l and a maximal value p_u for p, these integrals are easily computed for our simulations. Interestingly, because the large sample size conveys a lot of information to the consumer, the consumer is in the same situation as in the case of a price which does not change at all over time – at least for the a-heuristics. For the z-heuristics , the consumer's optimal choice depends on prices and one would not expect her to do as well in a world of changing prices as she would do with the a-heuristics.

For our second, possibly more relevant modeling of the case of stochastic prices, our consumer encounters a potentially different price p_t in each period t. She bases her decision on the change of the "indirect utility" ΔV which now depends not just on the variable $(a_{t-1}, \Delta_{t-1}a)$ but also on the changing prices. The definition of the change, following the definition in (11), is as follows:

$$\Delta V(a_{t-1}, \Delta_{t-1}a, p_{t-1}, p_{t-2}) = V(a_{t-1}, p_{t-1}) - V(a_{t-2}, p_{t-2}). \tag{33}$$

The counterpart of (12) reads accordingly.

Note that the consumer does not actually have to remember prices but only past levels of her well-being which are connected with the underlying parameters and prices.

While, at first glance, the process has lost its Markovian property, this is not necessarily the case because, for the Cobb-Douglas utility function, the prices enter in a particularly simple way. This allows the formulation and computation of the non-degenerate values of $P_{(\cdot,\cdot)}$ and $P^{(\cdot,\cdot)}$ in the matrices of transition probabilities. For the interested reader, we will sketch the derivation of the probabilities in the Appendix only. Because the simulation does not require the use of the transition matrix, the case of stochastic prices will be included in the next section.

6 Simulation Results

For larger sets $\mathcal{G}$ and $\mathcal{H}$, analytic results for the stable vectors r, respectively s can be obtained similarly but are cumbersome to present. We therefore rely on simulations when investigating less coarse grids $\mathcal{G}$ and $\mathcal{H}$, respectively. For $\alpha = .5$ and the periods $t = 1000$ and $t = 2000$, Tables 5.1 to 5.4 present the actual welfare losses

$$L(a_t) \quad \text{and} \quad M(z_t), \text{ respectively,} \tag{34}$$

as well as the average welfare losses

$$AL(a_t) \quad = \quad \frac{1}{t}\sum_{q=1}^{t} L(a_q) \ \text{ and} \tag{35}$$

$$AM(z_t) \quad = \quad \frac{1}{t}\sum_{q=1}^{t} M(z_q), \text{ respectively,} \tag{36}$$

for various ε, δ-constellations in the range $\varepsilon \leq .1$ and $1 - \delta \leq .1$.

Here Tables 5.1 and 5.2 illustrating the welfare effects of the a-dynamics are based on

$$\left\{ \begin{array}{cc} \text{grid size:} & g = 1/30 \text{ and} \\ \text{initial conditions:} & a_0 = g \text{ and } \Delta_0 a = g. \end{array} \right. \tag{37}$$

Tables 5.1 and 5.3 are computed for a deterministic price p which is constant over time.[5] For the four tables, τ denotes the actual welfare loss at the corresponding time period t while the symbol ϕ denotes the average welfare loss incurred up to that time period.

	δ	.9		.999		.99999	
ε		$t = 1000$	$t = 2000$	$t = 1000$	$t = 2000$	$t = 1000$	$t = 2000$
.1	τ	.0022	.0022	.0000	.0022	.0000	.0000
	ϕ	.0106	.0061	.0049	.0030	.0049	.0030
.00001	τ	.1156	.1156	.0000	.0000	.0000	.0000
	ϕ	.1181	.1168	.0041	.0021	.0049	.0030

Table 5.1: a-dynamics for a deterministic price p

[5]The actual price does not matter. The results do not depend on it.

For Tables 5.2 and 5.4, the stochastic prices are determined according to the uniform distribution over a narrow ($p_l = .9, p_u = 1.1$) and a wide ($p_l = .1, p_u = 1.9$) price range. The correponding rows in the tables are labeled n and w respectively.

ε		δ	.9		.999		.99999	
			$t = 1000$	$t = 2000$	$t = 1000$	$t = 2000$	$t = 1000$	$t = 2000$
.1	n	τ	.0022	.0022	.0022	.0089	.0000	.0089
		ϕ	.0118	.0129	.0134	.0108	.0128	.0099
	w	τ	.0022	.0022	.0089	.0835	.0089	.0835
		ϕ	.0646	.0609	.0699	.0627	.0874	.0734
.00001	n	τ	.0572	.0572	.0000	.0089	.0000	.0000
		ϕ	.0600	.0586	.0143	.0121	.0130	.0120
	w	τ	.5011	.5011	.0835	.0835	.4000	.1541
		ϕ	.5014	.5013	.0979	.0907	.0756	.0669

Table 5.2: a-dynamics for stochastic prices with narrow and wide range

Tables 5.3 and 5.4 for the z-dynamics are based on

$$\begin{cases} \text{grid size:} & h = 1/30 \text{ and} \\ \text{initial conditions:} & z_0 = h \text{ and } \Delta_0 z = h. \end{cases} \tag{38}$$

ε	δ	.9		.999		.99999	
		$t = 1000$	$t = 2000$	$t = 1000$	$t = 2000$	$t = 1000$	$t = 2000$
.1	τ	0.0022	0	0	0	0.0022	0.0022
	ϕ	0.0048	0.0031	0.0039	0.0025	0.0039	0.0025
.00001	τ	0.5011	0.5011	0.0022	0.0022	0.0022	0.0022
	ϕ	0.5013	0.5012	0.0044	0.0033	0.0039	0.0025

Table 5.3: z-dynamics for a deterministic price p

ε		δ	.9		.999		.99999	
			$t = 1000$	$t = 2000$	$t = 1000$	$t = 2000$	$t = 1000$	$t = 2000$
.1	n	τ	0.0002	0.0096	0.0387	0.0032	0.0002	0.0180
		ϕ	0.0192	0.0162	0.0165	0.0141	0.0203	0.0144
	w	τ	0.0098	0.1514	0.1692	0.0772	0.0634	0.0904
		ϕ	0.1709	0.1709	0.1486	0.1445	0.1695	0.1580
.00001	n	τ	0.1638	0.1980	0.0001	0.0025	0.0010	0.0136
		ϕ	0.1854	0.1845	0.0148	0.0113	0.0136	0.0130
	w	τ	0.6143	0.5266	0.6589	0.0020	0.0149	0.0007
		ϕ	0.5165	0.5127	0.1874	0.1835	0.2004	0.1902

Table 5.4: z-dynamics for stochastic prices with narrow and wide range

As can be seen from Tables 5.1 and 5.3, for a deterministic and constant price, actual welfare losses eventually are fairly constant for both kinds of dynamics. As has to be expected with simulations of stochastic processes, there are exceptions. See $\varepsilon = .1$ and $1 - \delta = .001$ in the case of the a-dynamics and $\varepsilon = .1 = 1 - \delta$ for the case of the z-dynamics. Otherwise the actual welfare loss is the same for $t = 1000$ and $t = 2000$ whereas the average welfare loss decreases with larger t since the initial inefficiencies (due to starting with the lowest value) average out over a larger number of periods. The exceptional result for $\varepsilon = .00001$ and $\delta = .9$ with a nearly constant major welfare loss over time for the z-dynamics illustrates how rare mutations and adaptation failure (due to $1 - \delta = .1$) can induce long stationary phases.

For the stochastic prices, the results in Tables 5.2 and 5.4 suggest that for the

- *narrow price range* the actual welfare loss for $t = 2000$ is often not smaller than for $t = 1000$ (exception: $\varepsilon = .1, 1 - \delta = .001$ in the case of the z-dynamics), whereas average welfare losses mostly decrease from $t = 1000$ to $t = 2000$ (exception: $\varepsilon = .1 = 1 - \delta$ for the a-dynamics);
- *wide price range* the actual welfare losses can increase and decrease (z-dynamics: 4 decreases, 2 increases, a-dynamics: 1 decrease, 2 increases), whereas average results generally improve over time with the only exception: $\varepsilon = .1 = 1 - \delta$ for the $z-$dynamics.

Let us say that a constellation looks *unambiguously promising* if $\tau = 0$ for both $t = 1000$ and $t = 2000$ and $\phi_{2000} < \phi_{1000}$ with the notation used in Tables 5.1 to 5.4. For the deterministic price, only the results for the a-dynamics with $(\varepsilon = .1, 1-\delta = .00001), (\varepsilon = .00001, 1-\delta = .001$ or $.00001)$ and for the z-dynamics with $\varepsilon = .1, 1 - \delta = .001$ appear in this sense unambiguously promising. This confirms the intuition that in case of the z-dynamics a robust inefficiency appears more likely. For stochastic prices, only the constellation $(\varepsilon = .00001 = 1 - \delta)$ for the a-dynamics for the narrow price range looks unambiguously promising. An unambiguous convergence to efficient consumption behavior is thus rather an exception than the rule in case of stochastic prices, as had to be expected. More generally this seems to justify the following

Remark: Directional (or in the terminology of Huck et al. (1998): *trial and error*) learning will rarely bring about efficiency within a reasonable time span of behavioral adaptation (after less than 2000 possible steps of adaptation). A stochastic environment will further hinder quick convergence to efficiency.

Let us look at the effects of reducing the mutation rate ε from .1 to .00001. As can be seen from Tables 5.1 to 5.4, there often is no decrease of the actual or of the average welfare loss, leading to the next

Remark: If mutations become rare, improving efficiency is usually delayed.

Increasing the adaptation rate δ from .9 via .999 to .99999 often but not always causes a reduction of actual (see, for instance, for $\varepsilon = .1$ and $t = 2000$ the change from $\delta = .9$ to $\delta = .999$ for the narrow price range in Table 5.2 or the corresponding result for $\delta = .999$ to $\delta = .99999$ in Table 5.4) or average (see, for instance, for $\varepsilon = .00001$ and $t = 2000$ the change from $\delta = .999$ to $\delta = .99999$ in Table 5.4) welfare losses. Especially the latter possibility that even average welfare losses increase are somewhat discouraging. This is stated in the last

Remark: Increasing the adaptation δ beyond .9 does not necessarily enhance efficiency.

7 Conclusion

In the situation analyzed above, an individual decision maker can adapt to her decision environment which can be a deterministic or stochastic framework. Although, knowing the true measure of satisfaction, the optimal decision behavior could in principle be easily derived, the consumer is unable to anticipate how satisfaction depends on behavior. Thus she has to experiment with behavioral changes when ex post it can be judged whether a behavioral change led to an improvement or not (in case of stochastic prices one can, of course, be misguided).

Here we were not primarily interested in analytic convergence results (that there exist asymptotically stable choices or limit cycles) which, in our finite setup, is obvious. Some results are given in Section 4. What we wanted to study is rather the adaptation process itself e.g. by investigating whether early adaptation ($t = 1$ to 1000) differs substantially from later one ($t = 1001$ to 2000) when limiting attention to a reasonable number (≤ 2000) of possible adaptation steps.

Although our final goal is to study the interaction of individually adapting agents, which will be included in Part II, it is important to learn first which welfare losses are caused by the stochastic adaptation alone, both for the deterministic and the stochastic price environment. In the light of such results, we will hopefully be able to single out what is due to the interaction of trading partners who might adjust behavior simultaneously as assumed by Huck et al. (1998), or alternatingly as for instance, assumed by Cournot (1838) in his dynamic justification of the duopoly solution.

Our main finding is a surprising persistency of welfare losses for a substantial time range ($t \leq 2000$) which, not unexpectedly, is much smaller for deterministic and constant prices than for stochastic prices. For the a-dynamics, the highest average inefficiency for $t = 2000$ is .1168 for the deterministic case and .5013 for the case of stochastic prices in Tables 5.1 and 5.2. Similarly, for the z-dynamics, it is .5012 in the deterministic case and .5127 in the case of stochastic prices in Tables 5.3 and 5.4. This clearly confirms the obvious intuition that in a stochastic environment learning will only slowly improve behavior.

While the individual adjustment is stochastically influenced by the parameters $0 < \varepsilon, \delta < 1$ and by possibly random prices, the interaction of trading partners introduces a new type of uncertainty that will be an essential aspect to be studied in Part II. Of course, in a general equilibrium framework prices cannot be assumed to be exogeneously given (deterministic or not) , but will have to be derived from market clearing. This does not necessarily imply that for given individual behavior prices are deterministic. To generate deterministic and stochastic prices in Part II, similar to the distinction in Part I, we will rely on deterministic and stochastic endowments.

Appendix

In determining the probabilities $P_{(a_{t-1},\Delta_{t-1}a)}$ and $P^{(a_{t-1},\Delta_{t-1}a)}$ respectively, the consumer compares her well-being during the two preceding periods. For the case of the a-heuristics, this comparison, with $a_{t-2} = a_{t-1} + \Delta_{t-1}a$, amounts to evaluating the inequality

$$\frac{a_{t-1}^{\alpha}(1-a_{t-1})^{1-\alpha}}{a_{t-2}^{\alpha}(1-a_{t-2})^{1-\alpha}} < \gamma \tag{39}$$

where $\gamma = 1$ in the case of fixed prices p and in the case of large random sampling and consumption during the preceding periods: In the first case $\gamma = p^{\alpha}/p^{\alpha}$, in the latter it equals $\int p^{-\alpha} f(p)dp / \int p^{-\alpha} f(p)dp$. In this case, a_{t-1} , a_{t-2}, and α clearly determine whether this inequality holds or not. Thus the quantities $P_{(\cdot,\cdot)}$ and $P^{(\cdot,\cdot)}$ are either 1 or 0 .

If prices vary only between the periodic adjustments but not during the consumption periods, we obtain $\gamma = p_{t-2}^{\alpha}/p_{t-1}^{\alpha}$ with p_{t-1}, p_{t-2} being the realisations of the random prices. Thus, $P_{(a_{t-1},\Delta_{t-1}a)} = \text{Prob}\{\, p_{t-2}/p_{t-1} \leq d\}$ with d being the left side of (39) raised to the power $1/\alpha$. Note that this number is independent of the random prices. These probabilities are well-defined for all distributions of prices. If the random prices are **uniformly** distributed between p_l and p_u, then is it not hard to compute that

$$P_{(a_{t-1},\Delta_{t-1}a)} = \begin{cases} \frac{1}{2}(p_u/d - p_l)(p_u - dp_l)/(p_u - p_l)^2 & \text{for } d \leq 1, \text{ and} \\ 1 - \frac{1}{2}(p_u - p_l/d)(dp_u - p_l)/(p_u - p_l)^2 & \text{for } d \geq 1. \end{cases}$$

$P^{(\cdot,\cdot)}$ would be computed similarly. Note that the number d and thus the probabilities depend on $(a_{t-1}, \Delta_{t-1}a)$. While these numbers would be easy to compute, we are not using them in our simulations because we are more interested in the intermediate ($1 \leq t \leq 2000$) behavior for large state spaces which can be treated without explicitly using the matrix of transition probabilities.

References

Bush, E., and F. Mosteller (1955), *Stochastic Models for Learning*, New York: Wiley.

Cournot, A. (1838), *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, Paris: L.Hachette.

Fudenberg, D., and D.K. Levine (1998), *The Theory of Learning in Games,* Cambridge and London: The MIT Press.

Hammerstein, P., and R. Selten (1994), "Game Theory and Evolutionary Biology," *Handbook of Game Theory, Volume 2*, Aumann, R.J., and S. Hart (eds.), Amsterdam: Elsevier Science B. V., Chapter 28, 931-993.

Huck, S., H.-T. Normann and J. Oechsler (1998), "Through Trial and Error to Collusion," *mimeo,* Humboldt-University Berlin.

Kemeny, J.G., and J.L. Snell (1960), *Finite Markov Chains,* New York and Berlin: Springer-Verlag.

Mas-Colell, A., M.D. Whinston, and J.R. Green (1993), *Microeconomic Theory,* New York and Oxford: Oxford University Press.

Roth, A.E., and Erev, I. (1998), "Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review* 88 (4), 848 -881.

Selten, R., and J. Buchta, (1999) "Experimental Sealed Bid First Price Auction with Direcly Observed Bid Functions," *Games and Human Behavior - Essays in Honor of Amnon Rapoport*, Budescu, D.V., I. Erev, and R. Zwick (eds.), Mahwah, NJ: Lawrence Erlbaum Associates.

Weibull, J.W. (1995), *Evolutionary Game Theory*, Cambridge and London: The MIT Press.